

Transformations between symmetric sets of quantum states

Vedran Dunjko

SUPA, Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK
Division of Molecular Biology, Ruđer Bošković Institute, Bijenička cesta 54, P.P. 180,
10002 Zagreb, Croatia

Erika Andersson

SUPA, Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK
E-mail: {vd51, e.andersson}@hw.ac.uk

Abstract.

We investigate probabilistic transformations of quantum states from a ‘source’ set to a ‘target’ set of states. Such transforms have many applications. They can be used for tasks which include state-dependent cloning or quantum state discrimination, and as interfaces between systems whose information encodings are not related by a unitary transform, such as continuous-variable systems and finite-dimensional systems. In a probabilistic transform, information may be lost or leaked, and we explain the concepts of leak and redundancy. Following this, we show how the analysis of probabilistic transforms significantly simplifies for symmetric source and target sets of states. In particular, we give a simple linear program which solves the task of finding optimal transforms, and a method of characterizing the introduced leak and redundancy in information-theoretic terms. Using the developed techniques, we analyse a class of transforms which convert coherent states with information encoded in their relative phase to symmetric qubit states. Each of these sets of states on their own appears in many well studied quantum information protocols. Finally, we suggest an asymptotic realization based on *quantum scissors*.

1. Introduction

Quantum information theory promises new and exciting ways to process information. Often the advantage a quantum protocol gives over a classical procedure lies in the fact that quantum states may be non-orthogonal. Classical information may be encoded in non-orthogonal quantum states, as is the case for example in quantum key distribution. The classical information then cannot be fully extracted from the quantum state alone. Many physical systems are candidates for the realization of quantum processing, and often they perform well at distinct tasks. Therefore, future quantum devices may well be hybrid systems with interfaces linking the different parts, just as our classical information processing devices are today. When classical information is encoded in a set of quantum states, the information encodings of one system may be incompatible with the encodings of another in a sense which has no classical analogue: the ‘source’ states may not be related to the corresponding ‘target’ states by a fixed unitary transformation. This occurs, for instance, when we consider transforms between states of systems of distinct dimensionalities such as typical qubit states and coherent or squeezed states of continuous-variable systems. When transferring information from one system to another where the encodings are incompatible in this sense, we must then either accept errors or resort to probabilistic scenarios where information may be lost, or leaked. This is important from an information-theoretic and cryptographic perspective. Information is no longer perfectly controlled by the emitting party.

Transformations that take a ‘source’ set of quantum states to a ‘target’ set of quantum states, where the states in the two sets are not pairwise related by a single unitary transform, also have other applications. State-dependent quantum cloning is one example [1]. Another related and well-studied family of such transforms solve the problem of amplifying coherent light, while keeping the coherent phase unaltered [2, 3, 4, 5, 6]. This problem is very important in classical and quantum communication tasks over larger distances, and is usually resolved by generating approximations of amplified coherent states. Optimal measurements for distinguishing between quantum states can also be seen as transforms taking some set of quantum states to mutually orthogonal states, followed by a measurement to distinguish the latter from each other. For so-called minimum-error measurements, pioneered by Holevo and Helstrom, the transforms are allowed to err, i.e. the declared output need not always be correct. Another tradition requires correctness but allows for result which declares that the transform (measurement) has failed, following the works of Ivanović, Dieks and Peres [7, 8, 9]. Such measurements are then called unambiguous [10, 11].

Here we focus on unambiguous transforms taking pure states to pure states. For this setting there exists a convenient framework based on the structures of the Gram matrices of ‘source’ and ‘target’ states, developed by Chefles, Jozsa and Winter [12, 13], which we will briefly present. In these works, the sets of source and targets states are general, and finding transforms for given sets of source and target states is complicated. However, it is known that for the problem of distinguishing quantum states which

comprise a symmetric set, a simpler treatment is possible [10, 14, 15]. As we will show, this restriction simplifies the theory for general probabilistic transforms as well. As an application of the theory we develop, we study the properties of converting a set of coherent states to qubit states. This is an important example of an ‘interspecies’ transform, as these two types of encodings frequently appear in quantum information processing tasks.

2. Preliminaries

Our problem of interest is stated as follows: given two sets of pure states A and B (called ‘source’ and ‘target’ sets, respectively) of finite size N ,

$$A = \{|a_i\rangle\}_{i=1}^N; \quad B = \{|b_i\rangle\}_{i=1}^N,$$

what are the properties of a transform \mathcal{T} , allowed by quantum mechanics, which performs $\mathcal{T}(|a_i\rangle) = |b_i\rangle$ for all i perfectly with a certain probability? The transform can fail to produce the desired output state, or succeed, and these two possible outcomes are reported, i.e. the transform is heralded.

In the most general case, the success probabilities may depend on which source state we start from. We then have the following statement:

Lemma 1 *There exists a probabilistic transform taking each state $|a_i\rangle$ in A to the state $|b_i\rangle$ in B , succeeding with the probabilities p_i , for $i = 1, \dots, N$, iff there exist Gram matrices of kets Π^s and Π^f such that the equality*

$$G_A = P^s \circ \Pi^s \circ G_B + P^f \circ \Pi^f \quad (1)$$

holds, where

$$P^s = \left[\sqrt{p_i p_j} \right]_{i,j} \quad \text{and} \quad P^f = \left[\sqrt{(1-p_i)(1-p_j)} \right]_{i,j}, \quad (2)$$

and G_A and G_B are the Gram matrices of sets A and B .

This is a special case of the Theorem 3 in [12]. In the Lemma above, \circ denotes the Hadamard (Shur, point-wise) matrix product, and the Gram matrix of the set of kets (or more generally vectors) $A = \{|a_i\rangle\}_{i=0}^{N-1}$ is given by

$$G_A = [\langle a_p | a_q \rangle]_{p,q}, \quad p, q = 0, \dots, N-1.$$

The necessary and sufficient conditions for a matrix M to be a Gram matrix of normalized kets (states) are that *i*) M is a positive-semidefinite matrix, and *ii*) M has unity across the main diagonal.

Such a quantum transform can be equivalently viewed, in the spirit of the Stinespring dilation, as a unitary transform acting on an augmented Hilbert space,

$$U|a_i\rangle|0\rangle|0\rangle = \sqrt{p_i}|b_i\rangle|\psi_i\rangle|0\rangle + \sqrt{1-p_i}|Fail\rangle|\phi_i\rangle|1\rangle \quad \text{for all } i, \quad (3)$$

where we learn whether the transform has succeeded or failed by measuring the third ‘indicator’ register on the right-hand side of the expression. One may show that the

matrices Π^s and Π^f in expression (2) are the Gram matrices of the sets of kets $\{|\psi_i\rangle\}_i$ and $\{|\phi_i\rangle\}_i$, respectively. If the transform in equation (3) succeeds, then the output registers contain the target state $|b_i\rangle$ but also a residual state $|\psi_i\rangle$ which may be correlated with the input state. From an information-theoretic perspective, this residual state may be seen as a leak of information, hence we call the set of states $\{|\psi_i\rangle\}_i$ the *leak*. If the states $|\psi_i\rangle$ are not correlated with the input state, which happens if and only if $|\psi_i\rangle = |\psi_j\rangle$ for all i and j , then the transform is called *leakless*. Analogously, in case the transform fails, a fixed fail state is produced along with a residual state $|\phi_i\rangle$ is produced. The residual state may be correlated with the input state, and may be used to subsequently attempt to reconstruct the desired outcome. For this reason we call the set of states $\{|\phi_i\rangle\}_i$ the *the redundancy*. If all the states in the redundancy are identical, only then is the residual state uncorrelated to the input state, and the transform is called *redundancy-free*.

If the success probabilities above do not depend on the source state ($p_i = p_j$ for all i, j), we call the transform *uniform*. For this case the notion of the optimal transform can be naturally defined: a uniform probabilistic transform is optimal, if no other transform succeeds with a strictly greater probability.

Deterministic and unitary transforms are easily seen to be special cases of probabilistic transforms. For a deterministic transform, it holds that $p_i = 1$, in which case the criterion reads $G_A = \Pi^s \circ G_B$ (for some Gram matrix of states Π^s). For a unitary transform the complex matrix Π^s is an outer product of a vector, containing roots of unity, with itself (c.f. [12]) \ddagger . Throughout this paper, with G_S we will denote the Gram matrix of the set of states S , and with λ_{G_S} a vector comprising the eigenvalues of the matrix G_S . With I we denote the identity matrix, and with $\mathbf{1}$ we denote the matrix with unity at each entry.

2.1. Example: uniform unambiguous discrimination of pure states

Unambiguous discrimination of states (UDS) identifies the input state from a pre-defined set of states, error free, but allows a ‘failure’ option. It is equivalent to a probabilistic transform for which the states $|b_i\rangle$ are mutually orthogonal. The criterion for the existence of such a transform is given by Lemma 1. Since the Gram matrix of orthogonal states is the identity, and the Hadamard product of the identity and Gram matrix of states is the identity again, for the special case where the success probability p is independent of the source state (uniform UDS), the existence condition simplifies to the inequality

$$G_A - pI \geq 0, \tag{4}$$

meaning that the matrix $G_A - pI$ is positive-semidefinite. Since unitary basis change preserves operator positivity, and G_A is positive-semidefinite, hence diagonalizable in

\ddagger This freedom in the complex phases reflects the fact that kets in general contain information about the physically irrelevant complex phase.

an orthonormal basis, this implies and is implied by

$$p \leq \min \lambda_{G_A},$$

where $\min \lambda_{G_A}$ denotes the smallest eigenvalue of the matrix G_A . From this condition we easily capture a known result: the optimal success probability of UDS is equal to the smallest eigenvalue of the Gram matrix G_A §. A consequence of this is another famous result: a set of states may be unambiguously discriminated if and only if that set of states is linearly independent. The latter is clear as the spectrum of G_A contains a zero element if and only if the set A is linearly dependent.

From the fact that unambiguous state discrimination is possible iff a set of states is linearly independent, it is easy to see that if a uniform probabilistic transform \mathcal{T} is optimal, then the redundancy is a linearly dependent set of states. To prove this, assume that a uniform probabilistic transform \mathcal{T} succeeds with probability p , and that the redundancy is linearly independent. Then, in the case of failure, one can run UDS on the redundancy, and if this succeeds (with probability $p' > 0$, due to linear independence), the target state can still be generated from the outcome. This overall procedure (\mathcal{T} followed by UDS in case of failure) comprises a uniform probabilistic transform \mathcal{T}' which performs the same task as \mathcal{T} but succeeds with probability $p' + p > p$. Hence \mathcal{T} could not have been optimal.

3. Transformations between symmetric sets of pure states

As noted, the case when the sets of states in focus is symmetric is of interest since many quantum protocols [16, 17, 19, 20, 18]) work with symmetric quantum states.

A set of (pure) states $A = \{|a_i\rangle\}_{i=0}^{N-1}$ is *symmetric* if there exists a fixed unitary U with the property

$$U|a_i\rangle = |a_{(i+1) \bmod N}\rangle \text{ for all } i.$$

The assumption that source and target states are symmetric allows us to link probabilistic and uniform probabilistic transforms. In this case, any probabilistic transform can be ‘uniformized’, as shown by the following lemma:

Lemma 2 (*Uniformization*) *If there exists a probabilistic transformation taking the states in A to states in B , which succeeds with the probabilities $\{p_i\}_{i=1}^N$, where A and B are symmetric sets of states, then there exists a uniform probabilistic transform taking the states in A to states in B which succeeds with probability*

$$p = \frac{1}{N} \sum_{i=1}^N p_i.$$

The proof of this lemma is given in the Appendix.

Additional properties of uniform transforms with symmetric source and target states are rooted in the structural properties of Gram matrices of sets of symmetric states:

§ This result was proven using different techniques and stated in a different formalism in [10].

Lemma 3 *A Gram matrix of kets is a circulant matrix if and only if the corresponding set of kets is symmetric.*

Proof of this lemma is given in the Appendix.

A circulant matrix is a square matrix, defined by its first row, for which the i^{th} row is the right-circular shift of the first row by $i - 1$ positions. Circulant matrices frequently appear in signal processing, and have two convenient properties: *i*) circulant matrices diagonalize when conjugated by the unitary discrete Fourier transform (DFT) matrix, and *ii*) the discrete Fourier transform of the first row of the circulant matrix is a vector containing the eigenvalues of the circulant matrix [21]. The discrete Fourier transform matrix of size N is the Vandermonde matrix of the N^{th} primitive roots of unity, given with

$$DFT = \left[\exp \frac{-2(p-1)(q-1)i\pi}{N} \right]_{p,q}, \quad p = 1, \dots, N, q = 1, \dots, N$$

which, when scaled by the pre-factor $1/\sqrt{N}$ becomes unitary, and which we then denote $uDFT$.

The criterion for the existence of a uniform probabilistic transform taking states from A to B , succeeding with probability p , is the existence of Gram matrices of states Π^s and Π^f such that the equation

$$G_A = p\Pi^s \circ G_B + (1-p)\Pi^f \quad (5)$$

holds. This is a slight simplification of the more general condition in Lemma 1.

In general, probabilistic uniform transforms with symmetric source and target sets may have leak and redundancy which are not symmetric. Nonetheless, the following lemma shows that such a transform has a variant with the same success probability where the leak and redundancy are symmetric:

Lemma 4 (*Symmetrization*) *If there exists a uniform probabilistic transform taking states from a symmetric set A to a set of symmetric states B , succeeding with some probability p , then there exists a uniform probabilistic transform taking the states from A to B , succeeding with probability p , where the leak and redundancy are symmetric.*

Proof of this lemma is given in the Appendix.

3.1. Finding optimal uniform transforms

Both from a practical and theoretical point of view, when considering transforms from a source to a target set one is often most interested in the optimal transforms. Optimality is naturally defined only in the case of uniform transforms. However, by virtue of Lemma 2, when transforms with symmetric source and target sets are concerned, if any kind of transform linking the source and target states exists, then so does a uniform transform. In this sense, for transforms between symmetric sets, optimality can in principle always be defined as the optimality of the uniform transform \parallel .

\parallel One may be tempted to do the same for non-symmetric transforms. However there exist non-uniform transforms which have non-symmetric source and/or target sets which succeed with non-zero probability

In general, given two sets of states A and B , the quest for the optimal uniform transform taking the states in A to states in B reduces to the maximization of the success probability p over the space of all positive-semidefinite matrices (of the appropriate size) Π^s and Π^f with unit diagonal, subject to the constraint given in expression (5). The dimensionality of the search space is then quadratic in the number of states. However, if source and target states are symmetric, as a consequence of Lemma 4, we may assume that Π^s and Π^f are circulant as well. Then all the matrices appearing in expression (5) are circulant, as the Hadamard product of circulant matrices is also circulant. Hence, they all diagonalize in the same basis, and the dimensionality of the search space reduces quadratically from $O(N^2)$ to $O(N)$, where N is the number of states.

The problem of finding optimal uniform transforms which have symmetric source and target sets is resolved by the following canonical linear program:

$$\begin{aligned} & \text{maximize } \vec{c}^T \cdot \vec{x} \\ & \text{subject to } M \cdot \vec{x} \leq \vec{b} \\ & \text{and } \vec{x} \geq 0, \end{aligned}$$

where $\vec{c}^T = [1, \dots, 1]$, $\vec{b} = \lambda_{G_A}$, and $M = DCM_{\lambda_{G_B}}$, which is a circulant matrix, where the i^{th} column is the vector λ_{G_B} ‘downward’ shifted by $i - 1$ positions (the *discrete convolution matrix* $DCM_{\lambda_{G_B}}$ of the vector λ_{G_B}). The optimal success probability is given by

$$p = \frac{\vec{c}^T \cdot \vec{x}}{N}, \quad (6)$$

where the dot ‘ \cdot ’ (e.g. $\vec{x}^T \cdot \vec{y}$ or $M \cdot \vec{x}$) denotes the standard matrix product. The vector of eigenvalues of the Gram matrix of the leak of the optimal transform is given by $\lambda_{\Pi^s} = \frac{1}{p} \vec{x}$. As both G_A and G_B are circulant matrices, the vectors of eigenvalues λ_{G_A} and λ_{G_B} are computed by taking the discrete Fourier transform of the first row of G_A and G_B , respectively.

In the remainder of this section we show that the linear program above solves the problem of finding optimal uniform transforms. The constraint (5) where all the matrices are circulant can be written in terms of the vectors of eigenvalues of the matrices appearing, as they all diagonalize in the same basis:

$$\lambda_{G_A} = p \lambda_{\Pi^s \circ G_B} + (1 - p) \lambda_{\Pi^f}. \quad (7)$$

Note that for the vector λ_{Π^f} to be a vector of eigenvalues of a circulant Gram matrix of states, it is sufficient and necessary that all its entries are non-negative and sum up to N . Using the circular convolution Theorem it can be shown that

$$\lambda_{\Pi^s \circ G_B} = \lambda_{\Pi^s} * \lambda_{G_B}, \quad (8)$$

where $*$ represents the (normalized) discrete convolution (or discrete cross-correlation) of vectors defined as follows. If \vec{x} and \vec{y} are two vectors of size N , with corresponding

for some states at least, for which no uniform transform exists (all uniform transforms fail with unit probability).

entries x_i and y_i for $i = 0, \dots, N-1$, then $\vec{z} = \vec{x} * \vec{y}$ is a length N (with components denoted z_i), defined component-wise by

$$z_i = \frac{1}{N} \sum_{j=0}^{N-1} x_j y_{[(N-j+i) \bmod N]}. \quad (9)$$

The discrete convolution of two vectors can also be represented in terms of a matrix-vector product by using the discrete convolution matrix $DCM_{\vec{x}}$ of the vector \vec{x} defined via its transpose: the transpose matrix $DCM_{\vec{x}}^T$ is a circulant matrix whose first row is the vector \vec{x} . It holds that $\vec{x} * \vec{y} = DCM_{\vec{x}}^T \cdot \vec{y} = DCM_{\vec{y}} \cdot \vec{x} = \vec{x} * \vec{y}$. Hence, the constraint (7) is equivalent to

$$\lambda_{G_A} = p DCM_{\lambda_{G_B}} \lambda_{\Pi^s} + (1-p) \lambda_{\Pi^f}, \quad (10)$$

which can be shown to be equivalent to the inequality

$$\lambda_{G_A} - p DCM_{\lambda_{G_B}} \lambda_{\Pi^s} \geq 0, \quad (11)$$

where λ_{Π^s} is a non-negative real vector, whose entries sum up to N . The inequality above is interpreted component-wise \P . To obtain the linear program stated at the beginning of this section, we note that if a vector \vec{x} is a vector of length N with non-negative entries $\{x_i\}$, which maximizes $s = \sum_{i=1}^N x_i$ subject to the constraint

$$DCM_{\lambda_{G_B}} \vec{x} \leq \lambda_{G_A}, \quad (12)$$

then $\lambda_{\Pi^s} = \frac{N}{s} \vec{x}$ allows for the maximal p subject to constraint (11), and the maximum is reached at $p = \frac{s}{N}$.

3.2. The geometric interpretation of the optimization procedure

As we have shown, the search for the optimal probability of success p_{opt} of a uniform transform which takes N input states to N output states, where both sets of states are symmetric, reduces to the following optimization problem:

p_{opt} is the maximal p subject to constraint

$$\lambda_{G_A} = p DCM_{\lambda_{G_B}} \lambda_{\Pi^s} + (1-p) \lambda_{\Pi^f}. \quad (13)$$

where λ_{Π^s} and λ_{Π^f} are some non-negative real vectors, whose entries sum up to N .

We have also shown that the constraint above is equivalent to the inequality

$$\lambda_{G_A} - p DCM_{\lambda_{G_B}} \lambda_{\Pi^s} \geq 0, \quad (14)$$

where λ_{Π^s} is a non-negative real vectors, whose entries sum up to N .

\P To prove this equivalence, note that (10) implies the constraint (11) as the eigenvalues of Π^f have to be non-negative. To see that the inverse holds as well, it suffices to show that if λ_{Π^s} is a vector of positive components which sum up to N , then the entries of the vector $\lambda_{G_A} - p DCM_{\lambda_{G_B}} \lambda_{\Pi^s}$ sum up to $(1-p)N$. By construction, the entries of λ_{G_A} sum up to N . Recall that $DCM_{\lambda_{G_B}} \lambda_{\Pi^s}$ is also a vector of eigenvalues of a Gram matrix of a symmetric set of kets. Hence its components also sum up to N . Hence, the components of $\lambda_{G_A} - p DCM_{\lambda_{G_B}} \lambda_{\Pi^s}$ sum up to $N - pN = (1-p)N$, and we have shown the equivalence of constraints (7), (10) and (11).

The search space defined by the constraint (11) is the space of all points embedded in an N dimensional space whose coordinates sum up to N . This is a convex set, defined by the extreme points $\{e_i\}_{i=1}^N$, where e_i is a vector with the number N as the i^{th} component, and zeroes elsewhere. But then, by the linearity of matrix-vector multiplication, the set

$$S = \left\{ DCM_{\lambda_{G_B}} \lambda_{\Pi^s} \mid \lambda_{\Pi^s} \geq 0, \|\lambda_{\Pi^s}\|_1 = N \right\}$$

is a convex set as well. The norm $\|\cdot\|_1$ is defined as the sum of the absolute values of the entries of the vector in the argument. It is easy to see that S is the convex hull of the columns of the matrix $N \times DCM_{\lambda_{G_B}}$. First, let us assume $DCM_{\lambda_{G_B}}$ is non-singular, which is equivalent to saying that the set of target kets does not contain mutually orthogonal kets. Then it holds that the columns of the matrix $N \times DCM_{\lambda_{G_B}}$ are also the extreme points of the set S . The constraint (11) can then be written as

$$\lambda_{G_A} - p X \geq 0 \tag{15}$$

where $X \in S$.

Let T be the set defined as follows:

$$T = \text{conv}(\{e_i\}_{i=1}^N),$$

where $\text{conv}(A)$ denotes the convex hull of the set of points A . T is the convex set of all points which correspond to a symmetric set of N states. This is a regular $(N-1)$ -simplex which we can embed in the vector space \mathbb{R}^N . Clearly, the point λ_{G_A} is an element of T , and S is also a regular $N-1$ -simplex, contained in T . Simplices T and S share their center at coordinates $(1, \dots, 1)$, and S is in a scaled down, rotated copy of T . From this it can be shown that the rows of the matrix $DCM_{\lambda_{G_B}}$ are the extreme points of the set S even when the discrete convolution matrix is singular. The only exception is the degenerate case when all the entries of $DCM_{\lambda_{G_B}}$ are equal, which corresponds to the case when the set of target states is an orthogonal basis.

It is easy to see that, if λ_{G_A} lies in S , then the constraint (11) can be satisfied for $p = 1$, *i.e.* there exists a deterministic transform from the set of states A to the set of states B . If this is not the case then the geometric interpretation of the constraint is as follows:

Lemma 5 *For a $0 < p \leq 1$ there exists a solution X satisfying the constraint (15) if the intersection between the simplex $p \times S = \{p \times \vec{x} \mid \vec{x} \in S\}$, embedded in \mathbb{R}^N , and the (N) orthotope (hyperrectangle or box) L defined by the opposite points $(0, \dots, 0)$ and λ_{G_A} is not the origin point alone.*

The orthotope L can be defined as

$$L = \left\{ \vec{x} \in \mathbb{R}^N \mid \lambda_{G_A} - \vec{x} \geq 0 \right\},$$

which makes the validity of the geometric interpretation above obvious. The geometric interpretation is illustrated in Figure 1, for the case $N = 3$.

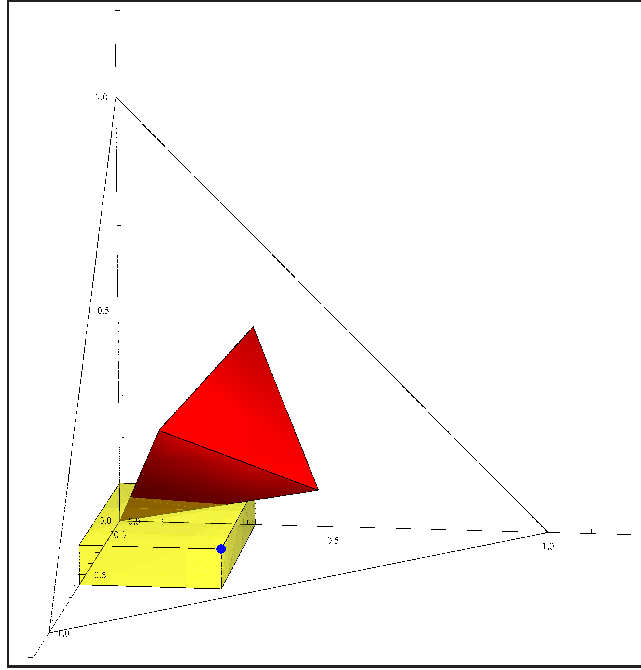


Figure 1. Illustration of the geometric interpretation of the solution existence criterion given in expression (15), for $N = 3$. The vector of eigenvalues of the Gram matrix of the source states is represented as a single point (blue in our illustration) which lies somewhere in the simplex defined by the extremal points $(0, 0, 1)$, $(0, 1, 0)$, $(1, 0, 0)$. This simplex is represented by the transparent triangle. The source states (represented by the blue point) uniquely define the orthotope L – a box, given in yellow. All points in the orthotope (and only those points) have the property that *i*) all their components are non-negative, and *ii*) any point in the orthotope when subtracted coordinate-wise from the blue point (defined by the source states) gives a point with non-negative components. The vector of eigenvalues of the Gram matrix of the target states defines the corresponding discrete convolution matrix, the columns of which are the extremal points of the search space S defined at the beginning of this section. The space S is, for $N = 3$, a regular 2-simplex (moreover, an equilateral triangle), embedded in a 3-dimensional space, and is represented with the red triangle which lies within the transparent triangle representing all possible vectors of eigenvalues of the Gram matrix of a size three symmetric set of states. Together with the origin, the extremal points of S define the 3-simplex $p \times S = \{p \times x | x \in S\}$ which is represented by the entire red tetrahedron in the illustration. Lemma 5 states that the necessary and sufficient criterion for the existence of a probabilistic transform taking the source to the target states is that the intersection between the red tetrahedron and the yellow box is not just the point of origin. The point of origin would correspond to a transform which succeeds with probability zero. Note also that clearly the intersection of the orthotope and the tetrahedron can be just the point of origin only if the orthotope has at least one dimension zero. This corresponds to the setting where source states are linearly dependent.

Let us now consider a few special cases, for illustration purposes. It is clear that if $S = T$ then for any set of input states the transform can be done deterministically, as $\lambda_{G_A} \in T$. However, if $S \neq T$, then $DCM_{\lambda_{G_B}}$ has exactly one ‘1’ in each row and each column, hence the vector of eigenvalues λ_{G_B} has one entry equal to N and the rest is zero. This corresponds to the setting where the target set of states comprises exactly one state, and the deterministic transformation performing this is the contraction to that particular state.

In the opposite scenario, S may consist of a single point – the point $(1, \dots, 1)$. In this case $DCM_{\lambda_{G_B}}$ is a matrix containing just unities, and the corresponding set of target states is then orthogonal. For there to exist a solution satisfying the constraint (15), by the geometric interpretation, the line $\{p(1, \dots, 1) | 0 < p \leq 1\}$ must intersect the orthotope L . This happens if and only if the extreme point λ_{G_A} which defines the orthotope has all components non-zero. This requirement implies that the input set of states is linearly independent. If we recall that a transformation with orthogonal target states is equivalent to unambiguous discrimination of input states, then we see we have recaptured a well-known result⁺: a set of states can be unambiguously discriminated if and only if the set is linearly independent.

Finally we can use the geometric interpretation to give a new result, which we haven’t addressed thus far: If a uniform transformation with symmetric sets of input and output states is optimal, then the leak is linearly dependent. To show this we will adopt a dynamic picture as illustrated in Figure 2. Let $\lambda_{G_A} \notin T$. What we seek is the largest p such that the simplex $p \times S = \{p \times \vec{x} | \vec{x} \in S\}$ and the orthotope L intersect. The simplex $p \times S$ clearly lies in the simplex $p \times T = \{p \times \vec{x} | \vec{x} \in T\}$, and the intersection will occur in this simplex. As we slowly decrease p the intersection between the simplex $p \times T$ and the orthotope L ‘grows’ while the simplex $p \times S$ slowly reduces in size. At one point, for some p , the intersection of the simplex $p \times T$ and the orthotope L touches the simplex $p \times S$, if there is a solution to the problem. Whenever this happens, the touching point is clearly on the face of the simplex $p \times S$, and not an interior point.

This means that the solution (the corresponding touching point in S) is a convex combination of at most $N - 1$ rows of $DCM_{\lambda_{G_B}}$, which in turn implies that λ_{Π^s} has a zero component. Since λ_{Π^s} is the vector of eigenvalues of the Gram matrix of the leak, this means the leak is linearly dependent. If we now join this with the fact that optimal transforms have a linearly dependant redundancy, shown in section 2.1, we get the following statement:

Lemma 6 *If a uniform transformation with symmetric sets of input and output states is optimal, then the leak and the redundancy are linearly dependent.*

The inverse however, does not hold.

⁺ Restricted, however, to symmetric sets of input states.

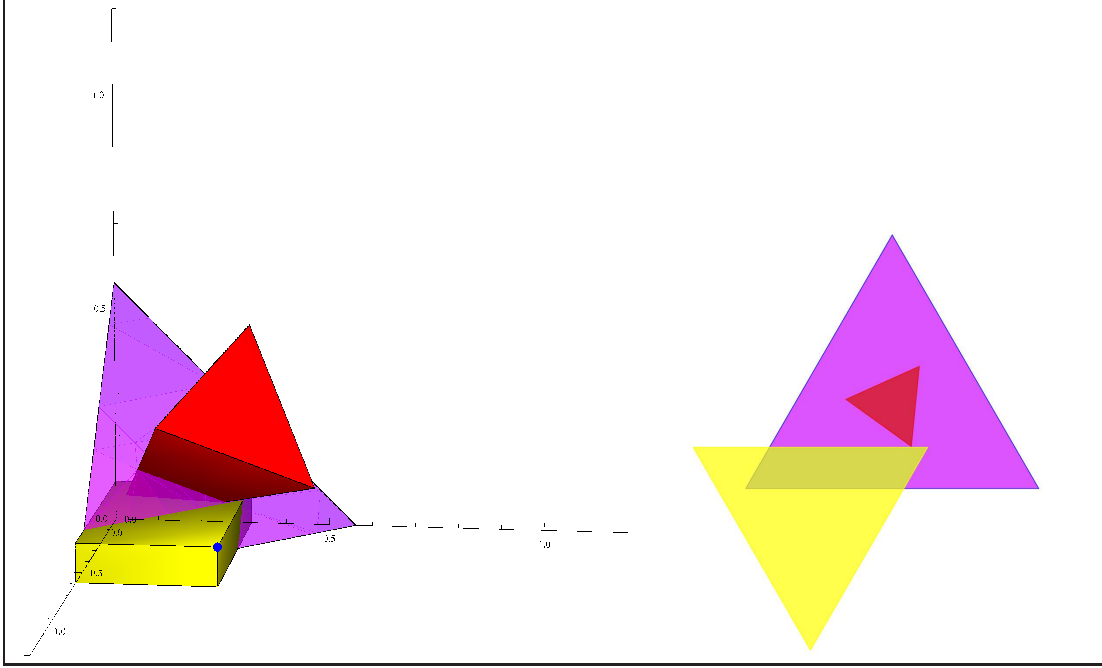


Figure 2. Dynamic picture: The left-hand side illustration represents the setting where the optimal transform has been found. For a particular (maximal) p the simplex which is the intersection of the (red) tetrahedron and the simplex $p \times T = \text{conv}((0, 0, p), (0, p, 0), (p, 0, 0))$ represented with the purple triangle touches the intersection of the yellow orthotope L and the same simplex $p \times T$. The right-hand side image illustrates this event when the view is restricted only to the $N - 1 = 2$ -dimensional unique hyperplane H_p containing the simplex $p \times T$. The intersection of the hyperplane H_p and the whole red tetrahedron makes up the red triangle, the purple triangle is the convex set $p \times T = \text{conv}((0, 0, p), (0, p, 0), (p, 0, 0))$ in the same hyperplane. The yellow triangle corresponds to the intersection of the hyperplane H_p and the extended orthotope L' where the edges of the original orthotope are allowed to extend to $-\infty$ each. This intersection is a regular 2-simplex again, with a fixed orientation with respect to the simplex $p \times T$. In the dynamic picture, if we were to let p slowly decrease from unity, we would witness the yellow triangle emerge from a single point, and grow until it touches the red triangle. The centre of the yellow triangle would slowly move towards the closest extremal point of $p \times T$ due to the change of its position in the barycentric coordinate system of the simplex $p \times T$ which changes as a function of p .

Geometric characterization of the leak and the redundancy

As we have shown, a uniform transform from symmetric to symmetric states succeeding with the probability p can always be realized in such a way that the leak and the redundancy are symmetric sets of states (Lemma 4). In this case, the leak and the redundancy can completely be characterized from the geometric picture. Recall that, if the transform exists for a fixed p , then the intersection between the simplex $p \times S$ and the orthotop L is non-empty and this intersection is contained in the simplex $p \times T$. It is easy to see that the intersection $p \times F = L \cap p \times T$ is a convex set, more precisely, a bounded convex polytope.

Let X be a solution, obeying the constraint (15). The vector X completely characterizes the leak. Recall, the vector X is of the form $DCM_{\lambda_{GB}} \lambda_{\Pi^s}$, where λ_{Π^s} is the vector of eigenvalues of the leak set. Thus, the vector X , viewed as a point in the simplex $p \times T$ embedded in the Euclidean space \mathbb{R}^N , is a convex combination of the rows of the matrix $DCM_{\lambda_{GB}}$. The weights of this convex combination are the components of λ_{Π^s} . In other words, the representation of X in the barycentric coordinates given by the extreme points of $p \times S$ (these points are the rows of the (scaled) matrix $DCM_{\lambda_{GB}}$) gives exactly the vector of eigenvalues of the Gram matrix Π^s . A barycentric coordinate system is a coordinate system in which a point's position is specified as the center of mass, or barycenter, of masses placed at the vertices of a simplex, in our case the simplex $p \times S$, which is the convex hull of the rows of the matrix $p \times DCM_{\lambda_{GB}}$.

An analogous observation can be done for the redundancy – X represented in the barycentric coordinates of some of the extreme points of $L \cap p \times T$ * will give us the structure of the redundancy. While this relationship is more involved than in the case of the leak, and we leave it for further research, certain easy observations can be made for the optimal transform case.

As we noted, if the transform is optimal, then the solution point X lies in the intersection of the faces of the polytope $p \times F = L \cap p \times T$ and the simplex $p \times S$, *i.e.* it is not in the interior of either. If the dimensionality of the face which is involved in the contact of $p \times S$ is zero (a vertex) then every symmetrized optimal transform is always leakless. Similarly, if the dimensionality of the face which is involved in the contact of $p \times F$ is zero, then it is redundancy-free.

If the contact involves faces of higher dimensionalities of $p \times F$ then essentially anything may happen, depending on the structure of the overlap. In the example given in the right-hand side illustration of Figure 2, the contact point for the simplex $p \times S$ (red) and the simplex $p \times L' \cap H_p$ (yellow) is a vertex of the simplex $p \times S$, and thus this transform is leakless. However, the contact point is interior of a 1-dimensional face of the yellow simplex, indicating that the vector of eigenvalues of the redundancy has two non-zero entries. Thus, the redundancy comprises at least two non-equal vectors.

Note that the structure of the overlap depends on the relative orientations and positions of the polytope $p \times F$ and the simplex $p \times S$. As we noted, the simplex $p \times S$ is just a scaled down and rotated simplex $p \times T$. The orientation of the polytope $p \times F$ is in a sense fixed with respect to the orientation of $p \times T$. To explain this, consider the simplex $L' \cap p \times T'$ where $L' = \prod_{i=1}^N \langle -\infty, \lambda_{GA}^i \rangle$, λ_{GA}^i being the i^{th} component of the vector λ_{GA} and the product is the Cartesian product. We define pT' to be the hyperplane defined by the points $\{p \times e_i\}_{i=1}^N$. The set L' is just the extended orthotope L where the sides (1-faces) radiating from the point λ_{GA} are allowed to stretch to $-\infty$. Then $L' \cap p \times T'$ is the intersection of $L' \cap p \times T'$ and the positive quadrant $\prod_{i=1}^N [0, \infty)$. $L' \cap p \times T'$ is then a regular N -simplex, and if we translate it by moving the center to the

* The number of the extreme point of this polytope may be larger than $N + 1$, but by Carathéodory's theorem, each point in this polytope can be represented as a convex combination of at most $N + 1$ points.

point (p, \dots, p) we have a simplex which is a scaled, centrally mirrored copy of $p \times T$. In this sense, the orientation of $p \times F$ (recall, $p \times F = (L' \cap p \times T') \cap \prod_{i=1}^N [0, \infty)$) is fixed, relative to the orientation of $p \times T$.

3.3. Quantifying the leak and the redundancy

Transformations between different types of quantum states become unavoidable when heterogeneous encodings are used for different aspects of quantum information tasks. In particular, such transform may be part of a cryptographic protocol, in which case quantifying the leak and redundancy in information-theoretic terms becomes crucial. For instance, one can imagine a simple two-party scheme in which party A, traditionally called Alice, wishes to send to party B, called Bob, information encoded in quantum states comprising the set of target states B . However, Alice has at her disposal only quantum states from a set of quantum states A . So, Alice indeed does send her information encoded as states in A to Bob, who performs an optimal probabilistic transform in order to obtain the target state B . For example, an ideal protocol may call for single-qubit states, but Alice can only generate pure states which approximate qubit states. It is then important for Alice to know what additional information Bob can obtain when transforming source states to target states \sharp . As we have seen, such a transform is characterized by an expression of the form $G_A = p\Pi^s \circ G_B + (1 - p)\Pi^f$ where the Gram matrices G_A and G_B fully characterize the source and target states (up to unitary equivalence), and Π^s and Π^f characterize the *leak* and the *redundancy*, that is, the residual states when the transform succeeds and when it fails, respectively. One way by which Alice may quantify the leak of information (embodied in the leak states) is by calculating the accessible information in this set of states. If $\{\rho_i\}_{i=0}^{N-1}$ is a set of quantum states, then the accessible information I_{acc} in this set of states is bounded above by the Holevo χ quantity, $I_{acc} \leq \chi(\rho_{AVG}) = S(\rho_{AVG})$, where $\rho_{AVG} = 1/N \sum_i \rho_i$ is the average state if each ρ_i appears equally likely as a message, $S(\cdot)$ denotes the Von Neumann entropy, and the last equality holds if ρ_i are pure. If $\lambda = (\lambda_0, \dots, \lambda_{N-1})$ is the vector of the eigenvalues of ρ_{AVG} , then the Von Neumann entropy can be expressed in terms of the Shannon entropy H as $S(\rho_{AVG}) = -\sum_{i=0}^{N-1} \lambda_i \log \lambda_i$.

If $A = \{|a_i\rangle\}_{i=1}^N$ is a set of kets (pure states), then using matrix algebra it can be shown that the non-zero eigenvalues of the matrix G_A and the operator $\sum_i |a_i\rangle\langle a_i|$ are equal. Hence, the upper bound on the accessible information in a set of states can be calculated as the Shannon entropy of normalized eigenvalues of the Gram matrix of that set of states. The optimization procedure we have presented, which finds the optimal success probability p , also finds the corresponding vector λ_{Π^s} . From this, λ_{Π^f} is easily computed, which are the eigenvalues of the Gram matrices of the leak and of the redundancy. From these eigenvalues it is then very simple to directly upper bound

\sharp Such approximations often appear in many proposals for realizations of quantum cryptographic protocols: polarization-encoded photons (which realize a qubit) are often approximated by polarized weak coherent pulses. In this case, almost without exception, a new security analysis is required.

the accessible information in the leak and the redundancy.

4. Application: From coherent states to qubit states

Traditionally, for most applications of quantum information processing, the information is encoded in qubit states. However, it is also possible to use continuous-variable states, that is, states of the quantum harmonic oscillator (e.g. coherent states). In this section the source states will be a set of coherent states

$$A = \{|a_k\rangle = |e^{i\theta_k}\alpha\rangle\}_{k=0,\dots,N-1} \quad (16)$$

where α is a real amplitude and θ_k are their phases. The target states are the qubit states in the Bloch sphere XY plane,

$$B = \left\{ |b_k\rangle = \frac{1}{\sqrt{2}} (|0\rangle + e^{i\theta_k}|1\rangle) \right\}_{k=0,\dots,N-1}. \quad (17)$$

By choosing the angles θ_k as $\theta_k = 2k\pi/N$ we obtain a very common family of encodings, which incidentally renders the sets A and B symmetric.

The problem we resolve is finding the optimal uniform transform taking the states in the set A to those in B . Initially, let us assume N is even. We may immediately note that the states in A are linearly independent, so an unambiguous measure-and-prepare process will get us the desired transform succeeding with the success probability of an UDS procedure applied on the states in A . The optimal success probability of such a UDS procedure establishes a lower bound, and an upper bound is found by noting that if N is even, then the desired probabilistic transform maps any two input states with relative phases differing by π into orthogonal states. Hence, in particular this transform effectively performs unambiguous discrimination of the states $|\alpha\rangle$ and $|\alpha\rangle$. By using the results of section 2.1, the success probability of this UDS procedure (hence of the overall probabilistic transform) is upper bounded by $s_{\text{bound}} = 1 - \exp(-2\alpha^2)$. This bound is always higher than the probability of unambiguous discrimination, except for the case of two states, where they coincide. The cases for 4 and 8 states are illustrated in Figure 3.

In the remainder of this section we prove, constructively, that the upper bound can always be reached. This is done by first obtaining results for the case $\alpha \leq 1$, and then using these results, constructing transforms also for the case $\alpha > 1$.

To begin, we introduce the notion of a multiprobabilistic transform, defined in [12]. Multiprobabilistic transforms are a generalization of probabilistic transforms, where there may be many different sets of targets states and with some probabilities an input state is transformed to a corresponding state in one of the target sets. For our purposes, we shall define the uniform version of such transforms:

Definition 7 Let $S = \{|s_i\rangle\}_{i=1}^n$ be a set of source states and $T^j = \{|t_i^j\rangle\}_{i=0}^n$ for $j = 0, \dots, k-1$ be a collection of possible target states. A uniform multiprobabilistic transform \mathcal{T} from the set S to the sets in $\{T^j\}_j$, succeeding with the probability vector

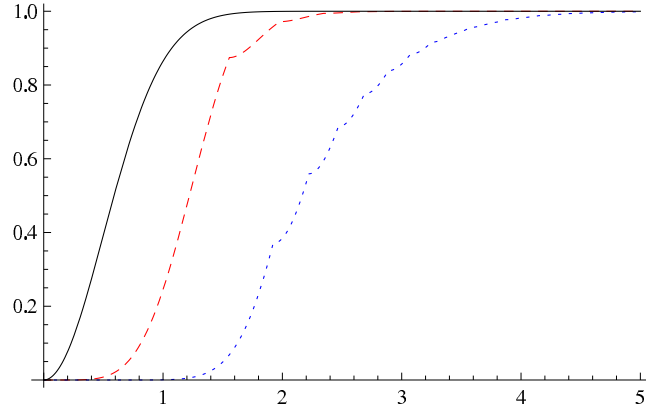


Figure 3. (color online) Comparison of the optimal success probability of unambiguous discrimination of 4 (red, dashed) and 8 (blue, dotted) states of a symmetric set of states, as a function of the real amplitude α . The black curve represents the optimal success probability of the coherent to qubit states transform, which is independent of the number of states.

(p_0, \dots, p_{k-1}) , where $\sum_{i=0}^{k-1} p_i = 1$ and for all i $p_i \geq 0$, performs

$$\mathcal{T}(|s_i\rangle) = |t_i^j\rangle \text{ with probability } p_j$$

for $i = 1, \dots, n$ and $j = 0, \dots, k-1$.

The set T^0 corresponding to success probability p_0 is reserved for the ‘fail outcome’ states, analogous to the redundancy set of states in probabilistic transforms.

As a consequence of Theorem 3 in [12], for fixed source set S and target sets $\{T^j\}_{j=1}^{k-1}$ and a probability vector (p_0, \dots, p_{k-1}) , such a uniform transform exists if and only if there exists a set of Gram matrices of states $\{\Pi^f, \Pi^1, \dots, \Pi^{k-1}\}$ such that the following equality holds:

$$G_S = p_0 \Pi^f + p_1 G_{T^1} \circ \Pi^1 + \dots + p_{k-1} G_{T^{k-1}} \circ \Pi^{k-1}, \quad (18)$$

where G_S is the Gram matrix of the set S and G_{T^j} the Gram matrix of the set T^j for all j . We will call such a transform leakless if the matrices $\Pi^j = \mathbf{1}$ for all j are matrices with all entries being the unity, and redundancy-free if the matrix $\Pi^f = \mathbf{1} \dagger\dagger$.

Let us now define a collection of sets of target states B^j as

$$B^j = \{|b_i\rangle^{\otimes j}\}_{i=0}^{N-1}, j = 1, 2, \dots, N-1. \quad (19)$$

That is, the set B^j comprises states which are j -fold copies of the elements of the (original target) set $B \equiv B^1$, which are the XY plane qubit states. Then we have the following lemma, which holds specifically for the source and target states of interest:

Lemma 8 *Let the amplitude α of the states in the set A , defined in equation (16), satisfy $0 < \alpha \leq 1$. Then there exists a uniform multiprobabilistic transform with the*

$\dagger\dagger$ Note that $\mathbf{1}$ is a Gram matrix of any set of unit vectors which are all equal.

success probability vector (p_0, \dots, p_{N-1}) , which takes the states from the set A to the collection of target states $\{B^j\}_{j=1}^{N-1}$ and is redundancy-free and leakless. The failure probability p_0 of this transform is equal to $\exp(-2\alpha^2)$.

The proof of this Lemma is somewhat cumbersome and left for the Appendix. The requirement that the transform be redundancy-free and leakless uniquely fixes the transform, up to the freedom in the choice of the realized fixed failure-outcome states.

As a corollary of this Lemma we obtain the desired uniform probabilistic transform from coherent states in A , for $0 < \alpha \leq 1$, to the qubit states in B , and a characterization of the leak and redundancy for this optimal transform, as we now show.

Corollary 1 *Let A and B be symmetric sets of an even number N states, as defined at the beginning of this section, and let $0 < \alpha \leq 1$. Then there exists a redundancy-free uniform probabilistic transform taking the states from A to corresponding states in B succeeding with probability $p_{\text{succ}} = 1 - \exp(-2\alpha^2)$. This transform is also optimal.*

Proof:

Lemma 8 establishes the existence of a multiprobabilistic uniform transform from the set A to the sets $\{B^j\}_{j=1}^{N-1}$, which is both redundancy-free and leakless, when $0 < \alpha \leq 1$. But then, by Theorem 3 in [12] there exists a probability vector (p_0, \dots, p_{N-1}) such that the following equality holds:

$$G_A = p_0 \mathbf{1} + p_1 G_{B^1} + \dots + p_{N-1} G_{B^{N-1}}. \quad (20)$$

Note, the expression above is the necessary and sufficient condition given in expression (18) for the existence of a uniform multiprobabilistic transform, which is now both redundancy-free, and leakless.

Since the Hadamard product is distributive, and by expression (A.11), this expression (20) can be rewritten as

$$G_A = p_0 \mathbf{1} + (1 - p_0) G_{B^1} \circ \left(\frac{p_1}{1 - p_0} G_{B^0} + \dots + \frac{p_{N-1}}{1 - p_0} G_{B^{N-2}} \right), \quad (21)$$

with $G_{B^0} = \mathbf{1}$. Let us denote expression in the parenthesis in the equation above by Π^s ,

$$\Pi^s = \frac{p_1}{1 - p_0} \mathbf{1} + \frac{p_2}{1 - p_0} G_{B^1} + \dots + \frac{p_{N-1}}{1 - p_0} G_{B^{N-2}}. \quad (22)$$

Note that Π^s is a Gram matrix of states, as it is a convex combination of Gram matrices of states. So we have

$$G_A = p_0 \mathbf{1} + (1 - p_0) G_{B^1} \circ \Pi^s. \quad (23)$$

This expression is a sufficient criterion for the existence of a uniform probabilistic transform taking the defined coherent states to qubit states. Since the fail probability is $p_0 = \exp(-2\alpha^2)$, by the upper bound on the success probability derived at the beginning of this section, it is the lowest possible, and this transform is optimal. This transform is also redundancy-free, as the Gram matrix of the redundancy is $\mathbf{1}$, that is, a Gram matrix of a set comprising identical states. The leak of this transform is symmetric by

lemma 3, as the matrix Π^s is a weighted sum of circulant matrices (see expression (22)), hence circulant itself. \square

By investigating the expression (22), we can construct the leak states of this transform explicitly. The leak state $|\psi_i\rangle$, corresponding to the input state $|a_i\rangle$ can, up to unitary equivalence, be written as

$$|\psi_i\rangle = \sum_{j=0}^{N-2} \sqrt{\frac{p_{j+1}}{1-p_0}} |b_i\rangle^{\otimes j} \otimes |0\rangle^{\otimes N-2-j} \otimes |j\rangle \quad (24)$$

where the states of the last register (the *indicator* register) are orthogonal for differing labels, and we define for any state $|\eta\rangle$, the zeroth tensoral power $|\eta\rangle^{\otimes 0} \equiv 1$ (the unity of the field underlying the Hilbert space, i.e. the number one). These ‘leaky’ states are superpositions of varying numbers of copies (from zero to $N-2$) of the target state $|b_i\rangle$, all living in orthogonal subspaces of a larger Hilbert space (due to the orthogonality of the indicator register states).

We will now prove the existence of an optimal transform for any amplitude, also $\alpha > 1$. To do this we first note that coherent states can be ‘split’ into multimode states of a lower amplitude, i.e. there exists an isometry performing $U|e^{i\phi}\alpha\rangle = \bigotimes_{k=0}^{M-1} |e^{i\phi}\beta_k\rangle, \forall \phi$, as long as $\alpha^2 = \sum_{k=0}^{M-1} \beta_k^2$. We note that in quantum optics, this transform can be implemented by using balanced beamsplitters and phase shifters. Assume that we are given a set of coherent symmetric states A , as defined in equation (16) with $\theta_k = 2\pi k/N$, of amplitude $\alpha > 1$. Each of these states in A can be deterministically taken to the state $\bigotimes_{k=0}^{M-1} |e^{i\theta_k}\beta\rangle$ by ‘splitting’ the coherent state into M modes, where $\beta = \frac{\alpha}{\lfloor \alpha \rfloor + 1}$ and $M = (\lfloor \alpha \rfloor + 1)^2$. Now we have that $\beta \leq 1$ and $\alpha^2 = M\beta^2$, where M is a non-negative integer. By the Corollary 1 we have that each subsystem state $|e^{i\theta_k}\beta\rangle$ can be individually transformed to the corresponding qubit state in the set B with probability $\exp(-2\beta^2)$. Note that, if only one of the individual transforms performed on the states $|e^{i\theta_k}\beta\rangle$ succeeds, then we have succeeded in generating exactly one copy of the target state from the source state $|e^{i\theta_k}\alpha\rangle$. The probability of the transform failing on all M copies is $\exp(-2\beta^2)^M = \exp(-2\alpha^2)$. Hence, we have the following Theorem.

Theorem 1 *Let A and B be symmetric sets of an even number N states, as defined in equation (16) with $\theta_k = 2\pi k/N$, and let $\alpha > 0$. Then there exists a redundancy-free uniform probabilistic transform taking the states in A to the corresponding states in B , succeeding with probability $p_{\text{succ}} = 1 - \exp(-2\alpha^2)$. This transform is optimal.*

The leak of this overall transform will in general comprise multimode states, which in some modes contain a fixed state (the modes where the probabilistic transform failed), and in some modes the target qubit and the individual transform leak of the form given in expression (24). In contrast to unambiguous discrimination procedures for symmetric coherent states, the success probability of these optimal transforms generating qubit states does not depend on the number of states. In this analysis, we have assumed that the number of possible input states is even. As the success probability does not depend on the (even) number of states, the probabilistic transform can be done with the same

success probability even when the number of states is N for an odd N . To see this, simply consider the transform which works for $2N$ states. The initial odd numbered symmetric states will be an interlaced subset of the extended set. However, here we do not have the validity of the upper bound any more, and it is not clear this success probability is optimal. While we do not offer a proof that the same bound holds for odd numbered states, evidence from performed numerical testing confirms this hypothesis.

An interesting aspect of the presented transform is that the success probability does not depend on the number of source and target states. Therefore it is possible that the same success probability may be reached when we consider the limit of an infinite number of states, $N \rightarrow \infty$. However, in the proofs of lemmas in this analysis, the finiteness of N is used, so proving this extension to the limit may be non-trivial. In the following section, we will however present a proposal for the realization of the presented transform, which does not assume a finite number of states, but achieves optimality in an asymptotic limit only.

4.1. Transforming coherent to qubit states using optical state truncation

After these results on the existence of optimal transforms, we will look at practical ways of implementing such transforms. A straightforward way of (sub-optimally) generating the desired qubit states from the source coherent states is through optical state truncation (OST) [22] or ‘quantum scissors’, as we will now describe. For a single mode state, such as a coherent state, OST is the probabilistic and heralded projection of the input state to a finite subspace (as defined by a selection of a number of Fock states), followed by renormalization of the state vector. OST has been realized using a linear optical network [23]. In this section we will focus on truncation to the subspace of the first two Fock states. Given the input state expanded in the number basis,

$$|\psi\rangle = \sum_{i=0}^{\infty} c_i |i\rangle,$$

where $\sum_i |c_i|^2 = 1$, OST is characterized by the POVM (POM) elements

$$\Pi_s = |0\rangle\langle 0| + |1\rangle\langle 1|, \quad \Pi_f = I - \Pi_s \quad (25)$$

and, upon success, produces the state

$$|\psi_{trunc}\rangle = \mathcal{N} (c_0 |0\rangle + c_1 |1\rangle)$$

where the normalization factor is $\mathcal{N} = (|c_0|^2 + |c_1|^2)^{-1/2}$. If we now consider the input state to be a state from our source set of N coherent states,

$$|a_j\rangle := |e^{\theta_j i} \alpha\rangle = e^{-\alpha^2/2} \sum_{k=0}^{\infty} \frac{\alpha^k e^{k\theta_j i}}{\sqrt{k!}} |k\rangle, \quad (26)$$

we see that the output state, after successful OST, which occurs with probability $p^{OST} = e^{-\alpha^2} (1 + \alpha^2)$, is

$$|a_j^{OST}\rangle = \frac{1}{\sqrt{1 + \alpha^2}} (|0\rangle + \alpha e^{i\theta_j} |1\rangle). \quad (27)$$

If $\alpha = 1$, this transform produces exactly the desired target qubit states.

This realisation does not, however, give the optimal success probability. The success probability of this transform for $\alpha = 1$ is approximately 0.735, which is less than the optimal value of approximately 0.864. The success probability of optical truncation to the vacuum and single photon subspace approaches unity more than exponentially quickly as the amplitude tends to zero. For $\alpha \neq 1$, the truncation will not produce the targeted qubit state, due to an uneven distribution of the weights between the $|0\rangle$ and $|1\rangle$ states. Re-weighting of the amplitudes can, however, also be achieved probabilistically, so now we consider the performance of the coherent to qubit transform realized by state truncation, followed by redistribution of the weights, for amplitudes $\alpha < 1$.

The redistribution of weights may optimally be done by applying a POVM defined by the positive elements

$$P_f = \gamma|0\rangle\langle 0|, \quad P_s = I - P_f, \quad (28)$$

where $\gamma = 1 - \alpha^2$. These transforms fall into a class we call *umbrella transforms*. The success rate (the probability of outcome associated with P_s) of this transform is $p_{umb} = 2\alpha^2/(1 + \alpha^2)$, hence the overall success probability of optical truncation followed by an umbrella transform for weight redistribution is

$$p_{overall} = p_{umb} p^{OST} = \frac{2\alpha^2}{1 + \alpha^2} e^{-\alpha^2} (1 + \alpha^2) = 2\alpha^2 e^{-\alpha^2}.$$

This value is always below the success probability of the optimal transform as the quotient $p_{opt}/p_{overall}$ is equal to

$$\frac{p_{opt}}{p_{overall}} = \frac{\sinh(\alpha^2)}{\alpha^2}, \quad (29)$$

which is always greater than 1 on the interval of interest, approaching unity when $\alpha \rightarrow 0$.

4.2. Asymptotic optimality through beamsplitting

While the optimal transform of coherent to qubit states cannot be realized by OST followed by an umbrella transform to redistribute relative weights, it is evident that this transform performs better and better as the amplitude is reduced. It is natural to check whether a beamsplitting pre-procedure, analogous to the one used to prove the optimality Theorem 1 in the $\alpha > 1$, may be used to boost the overall success probability.

The procedure goes as follows: the input state of real amplitude α is ‘beamsplitted’ into M modes of amplitude α/\sqrt{M} with the same complex phase as the initial beam (as was done in the proof of Theorem 1). Then OST is applied to each of the beams, and if an individual OST succeeds, an umbrella transform is applied to re-weigh the vacuum and $|1\rangle$ components. The overall procedure succeeds if, for at least one of the split off beams, both the truncation and the umbrella transform are successful.

As we have shown, for a real amplitude α , a re-weighted OST produces the corresponding qubit state succeeds with probability $p_{overall} = 2\alpha^2 e^{-\alpha^2}$. Then, the success

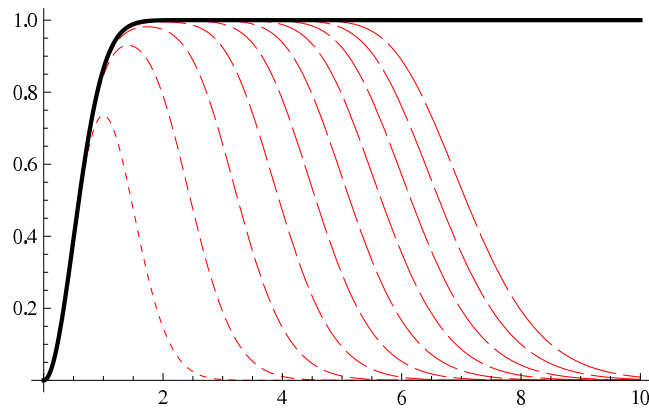


Figure 4. (color online) Comparison of the success probability of the optimal coherent to qubit states transform and the transform realized by beamsplitting into M beams of equal real amplitudes, followed by quantum scissors, followed by relative weight normalization between the vacuum and non-vacuum components on each of the weaker beams. The x axis gives the input amplitude α and y the success probabilities. The full (black) curve is the success probability of the optimal transform, and the (red) dashed curves the success probabilities of the beamsplitter-assisted quantum scissors strategies for $M = 1 \dots 10$. The longer-dashed curves correspond to larger parameter M .

probability of the strategy where the input beam has been split into M beams is given by

$$p_{\text{overall},M} = 1 - \left(1 - 2\frac{\alpha^2}{M}e^{-\alpha^2/M}\right)^M. \quad (30)$$

In the asymptotic case of infinite many ‘splits’, the failure probability becomes

$$p_{\text{overall},\infty} = \lim_{M \rightarrow \infty} \left(1 - 2\frac{\alpha^2}{M}e^{-\alpha^2/M}\right)^M = e^{-2\alpha^2},$$

which is equal to the failure probability of the optimal transform. The graph in Figure 4 compares the success probabilities of the optimal transform and the beamsplitter-assisted strategies for various numbers of splits M . This procedure can then arbitrarily well approach the optimal success probability. It is suitable for experimental realizations, as both quantum scissoring and the weight redistribution using umbrella transforms may be realized experimentally.

5. Conclusions

In this work we have addressed probabilistic transforms taking states from a ‘source’ to a ‘target’ set of quantum states, with emphasis on the case where these sets are symmetric. Such transforms can for example serve as interfaces between continuous-variable and finite-dimensional quantum systems. State-dependent cloning and quantum state discrimination are also special cases of probabilistic transforms.

We have emphasised that in a probabilistic transform, information may be lost and leaked, which may have impact on the protocol efficiency or security. For this purpose

we introduced the concepts of the leak and redundancy of a probabilistic transform. We have demonstrated how symmetric source and targets sets, which arise naturally in many quantum information applications, allows for a much simpler theory. In particular, we derived a linear program which finds optimal uniform probabilistic transforms in this symmetric setting. This constitutes a significant simplification over optimization techniques which must be employed in more general cases, and the dimensionality of the search space is reduced quadratically in the number of states considered. The presented method also allows for a simple characterization of the aforementioned leak and redundancy.

Following this, we applied the derived theory to the problem of transforming a particular set of coherent states to a particular set of qubit states. Both sets appear in many quantum information protocols. The considered set of coherent states are so-called ‘phase-locked’ quantum states (e.g. used for quantum key distribution) suitable for long-range communication, and the set of qubit states is ubiquitous in quantum computation. For this setting, we derived the optimal transform and characterized the leak and the redundancy. By using beamsplitting, followed by the well-studied process of optical state truncation or ‘quantum scissors’, and an experimentally feasible amplitude re-weighting procedure, a probabilistic transform between these sets of states can be realized, albeit with sub-optimal success probability. The success probability of this procedure can however be made to asymptotically approach the optimal success probability.

An immediate application of such a transform may be in the realization of Universal Blind Quantum Computation (UBQC) [19], in a case where the client is restricted to producing coherent states, in contrast to the single-qubit states required by the original protocol. A related procedure for this scenario was recently suggested in [24], where phase-randomized weak coherent states were used, and the information was encoded in the polarization. This encoding the information remained essentially unitarily equivalent to the original single-qubit encoding. The question whether UBQC is possible when the client uses phase-encoded coherent states (where the unitary equivalence no longer holds) remains open. Finally, the approaches developed in this paper may be applied to the task of amplifying coherent states *truly perfectly*, which can be achieved probabilistically when the number of possible phases is finite. This will be the subject of further work.

6. Acknowledgements

We gratefully acknowledge financial support by EPSRC grant EP/G009821/1.

Appendix A. Proofs of Lemmas

Here we give the proofs of the Lemmas which were stated in the main body of the paper. For the reasons of brevity, occasionally we will skip through technical details, and rather present the main ideas. We begin by proving the uniformization (Lemma 2)

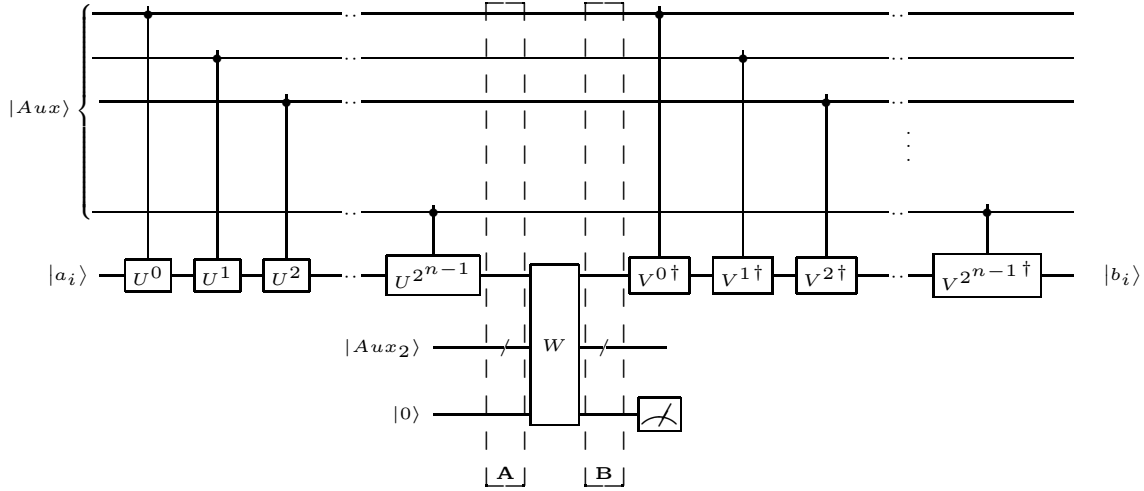


Figure A1. The quantum circuit by which a probabilistic transform, realized by the action of a unitary W acting on an augmented Hilbert space followed by the measurement of an indicator register, can be ‘uniformized’. The same circuit also serves to symmetrize the leak and the redundancy of a uniform probabilistic transform. In the proofs of lemmas we will address the states of the system above at cuts **A** and **B** denoted in this figure.

and symmetrization (Lemma 4) lemmas.

Proof of Lemmas 2 and 4

Lemma 2 (*Uniformization*) *If there exists a probabilistic transformation \mathcal{T} taking the states in A to states in B which succeeds with the probabilities $\{p_i\}_{i=1}^N$, where A and B are symmetric sets of states, then there exists a uniform probabilistic transform \mathcal{T}' taking the states in A to states in B which succeeds with probability*

$$p = \frac{1}{N} \sum_{i=1}^N p_i.$$

Proof:

As noted, each probabilistic transform may be realized as a unitary transform acting on an augmented Hilbert space, followed by a measurement of an indicator register. In the circuit of Figure A1, the transform \mathcal{T} is represented by this extended unitary W . As both the input and output sets of states are symmetric, there exist unitaries which sequentially shift through the states of the set, obeying the intrinsic order. We denote these unitaries by U and V , corresponding to the sets A and B respectively, and the controlled powers of these unitaries appear in the circuit. The state $|Aux\rangle$ is pre-set to

be the uniform superposition

$$|Aux\rangle = 1/\sqrt{N} \sum_{k=0}^{N-1} |k\rangle, \quad (\text{A.1})$$

where $|k\rangle$ is the l qubit state of the computational basis $|b_{l-1}\rangle \otimes \cdots \otimes |b_0\rangle$, $b_j \in \{0, 1\}$ for all j such that $(b_{l-1} \dots b_0)_2 = (k)_{10}$, where the subscripts designate the base of the number representations.

First let us show that the circuit shown performs the desired transform. The state of the system at cut **A** in the circuit is

$$1/\sqrt{N} \sum_{k=0}^{N-1} |k\rangle |a_{i+k \bmod N}\rangle |Aux_2\rangle |0\rangle, \quad (\text{A.2})$$

where $|Aux_2\rangle$ is some fixed auxiliary state in a sufficiently dimensional state space. The notation we shall use corresponds to the notation used in formula 3. Following this, the transform \mathcal{T} is applied to the register which contained the input state. The transform is explicitly realized as a unitary W acting on a bigger space. The state at cut **B** in the circuit is

$$\begin{aligned} & 1/\sqrt{N} \left(\sum_{k=0}^{N-1} \sqrt{p_{i+k \bmod N}} |k\rangle |b_{i+k \bmod N}\rangle |\psi_{i+k \bmod N}\rangle \right) |0\rangle \\ & + 1/\sqrt{N} \left(\sum_{k=0}^{N-1} \sqrt{1 - p_{i+k \bmod N}} |k\rangle |Fail\rangle |\phi_{i+k \bmod N}\rangle \right) |1\rangle. \end{aligned} \quad (\text{A.3})$$

If the measurement outcome of the indicator (the last) register corresponds to the state $|0\rangle$, then the transform has succeeded (c.f. expression 3). From the expression above, it can be seen that this happens with probability $p = \frac{1}{N} \sum_{i=1}^N p_i$. Assume that the indicator measurement yielded the desired output. The section of the circuit after cut **B** undoes the controlled rotations, and the state at the end of the entire circuit is

$$\mathcal{N} \sum_{k=0}^{N-1} \sqrt{p_{i+k \bmod N}} |k\rangle |b_i\rangle |\psi_{i+k \bmod N}\rangle. \quad (\text{A.4})$$

The middle register contains the desired output state, and the rest of the system contains a new leak. This overall procedure constitutes the new, uniformized probabilistic transform \mathcal{T}' from the statement of the Lemma, which succeeds with the averaged probability p . This proves Lemma 2. \square

One can verify that the new leak, generated by the ‘uniformized’ transform described above, comprises a symmetric set of states. Using an analogous analysis, one can show that the redundancy (state generated in case of the measurement outcome corresponding to the $|1\rangle$ state in the indicator register) is a symmetric set as well. Now, if the extended unitary W corresponds to a uniform probabilistic transform, with leak and redundancy which are not symmetric, then the extended transform of Figure A1 will have the same success probability as W itself, and the leak and redundancy will be symmetrized. Thus, the analysis above proves Lemma 4 as well. \square

Lemma 3 *A Gram matrix of kets is a circulant matrix if and only if the corresponding set of kets is symmetric.*

Proof:

Let $A = \{|a_k\rangle\}_{k=0}^{N-1}$ be a set of kets. We first show the necessity. If the set of kets is symmetric, then its Gram matrix is circulant. Let U be the unitary which sequentially shifts through the set of kets, obeying the intrinsic order. Then the Gram matrix may be written as

$$\begin{aligned} G_A &= [\langle a_p | a_q \rangle]_{p=0, q=0}^{N-1, N-1} = [\langle a_0 | U^{\dagger p} U^q | a_0 \rangle]_{p=0, q=0}^{N-1, N-1} = \\ &= [\langle a_0 | U^{q-p} | a_0 \rangle]_{p=0, q=0}^{N-1, N-1} = [\langle a_0 | U^{q-p \bmod N} | a_0 \rangle]_{p=0, q=0}^{N-1, N-1}. \end{aligned} \quad (\text{A.5})$$

It is easy to verify that the last matrix in the sequence of equalities above is circulant.

Next we show the sufficiency. If A is a set of states such that its Gram matrix G_A is circulant, then it is symmetric. Since G_A is a matrix of states, we have that G_A allows the Cholesky decomposition, that its spectrum $\{\lambda_k\}_{k=0}^{N-1}$ is real, non-negative and sums up to N (as the trace is preserved under basis change), and as it is circulant, we have that it diagonalizes in the $uDFT$ basis. Using these properties and a bit of matrix algebra, one can show that if a set of kets $\{|\psi_k\rangle\}_{k=0}^{N-1}$ has G_A as a Gram matrix, then its elements can be written as

$$|\psi_k\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \frac{1}{\sqrt{\lambda_j}} e^{\frac{2kj\pi i}{N}} |b_j\rangle, \quad (\text{A.6})$$

where the kets $\{|b_k\rangle\}_{k=0}^{N-1}$ comprise an orthonormal basis, and we define the coefficient $\frac{1}{\sqrt{\lambda_j}}$ to be zero if $\lambda_j = 0$. Consider the unitary U , acting on the $\{|b_j\rangle\}_j$ basis as follows:

$$U|b_j\rangle = e^{\frac{2j\pi i}{N}} |b_j\rangle. \quad (\text{A.7})$$

By applying U on the ket $|\psi_k\rangle$ we have:

$$\begin{aligned} U|\psi_k\rangle &= \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \frac{1}{\sqrt{\lambda_j}} e^{\frac{2kj\pi i}{N}} U|b_j\rangle = \\ &= \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \frac{1}{\sqrt{\lambda_j}} e^{\frac{2(k+1)j\pi i}{N}} |b_j\rangle = |\psi_{k+1 \bmod N}\rangle. \end{aligned} \quad (\text{A.8})$$

Hence, the set of kets A which they represent is symmetric and this proves the lemma.

□

The following lemmas were given in section 4. In their proofs we shall adhere to the notation of the main body of the paper.

Lemma 8 *Let the amplitude α of the states in the set A , defined in equation (16), satisfy $0 < \alpha \leq 1$. Then there exists a uniform multiprobabilistic transform with the success probability vector (p_0, \dots, p_{N-1}) , which takes the states from the set A to the collection of target states $\{B^j\}_{j=1}^{N-1}$ and is redundancy-free and leakless. The failure probability p_0 of this transform is equal to $\exp(-2\alpha^2)$.*

Proof:

As noted in the main body of the paper, the desired transform exists if and only if

$$G_A = p_0 \Pi^f + p_1 G_{B^1} \circ \Pi^1 + \cdots + p_{N-1} G_{B^{N-1}} \circ \Pi^{N-1} \quad (\text{A.9})$$

holds for a vector of probabilities (p_0, \dots, p_{N-1}) and for a set of Gram matrices of states $\{\Pi^f, \Pi^1, \dots, \Pi^{N-1}\}$. Acknowledging the requirement that this transform is leakless and redundancy-free the criterion becomes

$$G_A = p_0 \mathbf{1} + p_1 G_{B^1} + \cdots + p_{N-1} G_{B^{N-1}}, \quad (\text{A.10})$$

where $\mathbf{1}$ is a matrix with all entries being the unity.

The matrix G_{B^j} can be written as

$$G_{B^j} = \underbrace{G_B \circ \cdots \circ G_B}_{j \text{ times}} := G_B^{\circ j}, \quad (\text{A.11})$$

and since G_A , G_B are circulant, and the Hadamard product of circulant matrices is circulant, G_{B^j} is circulant for all j . Hence all the matrices in expression (A.10) simultaneously diagonalize in the unitary discrete Fourier transform basis so we can write this criterion in terms of vectors of eigenvalues of the corresponding matrices:

$$\lambda_{G_A} = p_0 \lambda_{\mathbf{1}} + p_1 \lambda_{G_{B^1}} + \cdots + p_{N-1} \lambda_{G_{B^{N-1}}}. \quad (\text{A.12})$$

The vector $\lambda_{\mathbf{1}}$ is the first vector of the canonical basis, that is vector with one as the first entry and zeroes elsewhere, multiplied by N .

It can be shown that, for any N , the vector of eigenvalues of G_B has only the first two eigenvalues non-zero, and their value is $N/2$. From this, using the properties given in expressions (8) and (9), we can see that, for $k \leq N-1$, the vector of eigenvalues of $G_B^{\circ k}$ is given by

$$\lambda_{B^k} = \frac{N}{2^k} \left[\binom{k}{0}, \binom{k}{1}, \dots, \binom{k}{k}, 0, \dots, 0 \right]^T. \quad (\text{A.13})$$

Let M be the column matrix defined by

$$M = [N e_1 | \lambda_B | \lambda_{B^2} | \cdots | \lambda_{B^{N-1}}]. \quad (\text{A.14})$$

Then we can rewrite the condition (A.12) as a system of equations,

$$\lambda_{G_A} = M \vec{p} \quad (\text{A.15})$$

where $\vec{p} = [p_0, \dots, p_{N-1}]^T$. Since M is upper-triangular, with non-zero element across the diagonal, it is invertible. Hence, there exists a unique vector \vec{p} satisfying the system above. The sum of the elements of a column of the matrix M is N , so we can see (by multiplying the system (A.15) with the row vector $\frac{1}{N} [1, \dots, 1]$ from the left) that $\sum_{i=0}^{N-1} p_i = 1$, as the sum of the eigenvalues of G_A is N .

To prove the stated Lemma, we need to show that all the values p_i are non-negative (for $0 < \alpha \leq 1$), and that we need to show that $p_0 = \exp(-2\alpha^2)$. We begin by showing the positivity of values p_i , as stated, and we finish of the proof by showing that $p_0 = \exp(-2\alpha^2)$.

As noted above, the system (A.15) has a unique solution (and $M^{(-1)}$ exists), and we need to show that the solution vector comprises positive elements, i.e.

$$M^{(-1)}\lambda_{G_A} \quad (\text{A.16})$$

is a vector of non-negative real numbers. Note that the matrix M can be written as $M = M'.D$, where M' collects all the binomial coefficients and D is a diagonal matrix which appropriately assigns the weights to the columns of M . The k^{th} column of matrix M' is then given by

$$\left[\binom{k}{0}, \binom{k}{1}, \dots, \binom{k}{k}, 0, \dots, 0 \right]^T.$$

The inverse of M is then

$$M^{(-1)} = D^{-1}.M'^{(-1)}. \quad (\text{A.17})$$

As the matrix D^{-1} comprises only positive elements (moreover it is also diagonal), in order to show that the expression (A.16) is a non-negative vector, it will suffice to show that

$$M'^{(-1)}\lambda_{G_A} \quad (\text{A.18})$$

is a non-negative vector. Let S be a diagonal matrix of size N of alternating signs, the first sign being positive. Using known properties of sums of binomial coefficients, one can show that $S.M'.S$ is the inverse of the matrix M' . We omit the proof of this claim as the proof is technical, and the details are of no further consequence.

Now we proceed to show that each entry of the vector

$$M'^{(-1)}\lambda_{G_A} = S.M'.S\lambda_{G_A} \quad (\text{A.19})$$

is non-negative, if the amplitude α is a positive and less or equal to unity. Let λ_i be the i^{th} eigenvalue of the matrix G_A , i.e. the i^{th} component of λ_{G_A} . Note that the enumeration starts at zero. Then the k^{th} entry of the vector $S.M'.S\lambda_{G_A}$ is given by

$$(e_k)^T S.M'.S\lambda_{G_A} = \sum_{j=k}^{N-1} (-1)^{j+k} \binom{j}{k} \lambda_j \quad (\text{A.20})$$

The last entry of the vector $S.M'.S\lambda_{G_A}$ is the last eigenvalue of G_A , hence positive, so for the expression (A.20) to be positive, it suffices to show that

$$\binom{j}{k} \lambda_j - \binom{j+1}{k} \lambda_{j+1} \geq 0 \quad (\text{A.21})$$

for all $0 \leq k < N-1$ and $k \leq j < N-1$. This expression simplifies to

$$\binom{j}{k} \lambda_j - \binom{j+1}{k} \lambda_{j+1} = \binom{j}{k} \left(\lambda_j - \frac{j+1}{j-k+1} \lambda_{j+1} \right). \quad (\text{A.22})$$

Since $\binom{j}{k}$ is positive, we only need to show that the following holds:

$$\lambda_j - \frac{j+1}{j-k+1} \lambda_{j+1} \geq 0. \quad (\text{A.23})$$

In order to show this, we need to analyse the structure of the eigenvalues appearing as components of λ_{G_A} . Recall, λ_{G_A} was defined as the discrete Fourier transform of the first row of G_A . Using the expansion of coherent states in the Fock basis the j^{th} eigenvalue can be given as

$$\lambda_j = \sum_{l=0}^{N-1} \exp(-2jl\pi i/N) \sum_{r=0}^{\infty} e^{-\alpha^2} \frac{\alpha^{2r}}{r!} \exp(2lr\pi i/n). \quad (\text{A.24})$$

This can further be rearranged as follows:

$$\lambda_j = e^{-\alpha^2} \sum_{l=0}^{N-1} \sum_{r=0}^{\infty} \exp(-2jl\pi i/N) \frac{\alpha^{2r}}{r!} \exp(2lr\pi i/n) \quad (\text{A.25})$$

$$= e^{-\alpha^2} \sum_{r=0}^{\infty} \frac{\alpha^{2r}}{r!} \sum_{l=0}^{N-1} \exp(2l(r-j)\pi i/n), \quad (\text{A.26})$$

where in order to get to expression (A.26), we used the fact that the infinite sum above is absolutely convergent, hence allows the commuting of sums.

By the properties of sums of roots of unity, the expression $\sum_{l=0}^{N-1} \exp(2l(r-j)\pi i/n)$ is equal to n if $r-j$ is divisible by N and zero otherwise. Hence we get

$$\lambda_j = e^{-\alpha^2} N \sum_{r=0}^{\infty} \frac{\alpha^{2(Nr+j)}}{(Nr+j)!}. \quad (\text{A.27})$$

The elements in the sum above appear as the summands in the Taylor expansion of $e^{2\alpha}$; for $j=0$, this sum collects every N^{th} summand from the Taylor series expansion, starting from the zeroth summand. For any other j it collects every N^{th} summand from the Taylor series expansion, starting from the j -th summand. We note that the eigenvalues above, for a fixed N can be expressed in a closed form in terms of Generalized hypergeometric functions.

We set out to show that inequality (A.23) holds. By inserting the explicit expressions for the eigenvalues we have derived, we obtain the expression

$$\lambda_j - \frac{j+1}{j-k+1} \lambda_{j+1} = e^{-\alpha^2} N \left(\sum_{r=0}^{\infty} \frac{\alpha^{2(Nr+j)}}{(Nr+j)!} - \frac{j+1}{j-k+1} \sum_{r=0}^{\infty} \frac{\alpha^{2(Nr+j+1)}}{(Nr+j+1)!} \right), \quad (\text{A.28})$$

and again by absolute convergence of the sums above we may reshuffle them and obtain

$$e^{-\alpha^2} N \sum_{r=0}^{\infty} \frac{\alpha^{2(Nr+j)}}{(Nr+j)!} \left(1 - \frac{j+1}{(j-k+1)(Nr+j+1)} \alpha^2 \right). \quad (\text{A.29})$$

The expression above is positive if the expression in the last parenthesis is positive. Now we inspect the coefficient with the term α^2 in the parenthesis,

$$\frac{j+1}{(j-k+1)(Nr+j+1)}.$$

This expression is always positive, and note that the denominator $(j - k + 1)$ is greater or equal to unity, and the denominator $(Nr + j + 1)$ is larger or equal to $j + 1$, so the entire expression is less or equal to unity. But then for $\alpha \leq 1$ the expression (A.29) is non-negative.

To finish the proof we need to show that $p_0 = \exp(-2\alpha^2)$. Note that $p_0 = e_0^T M^{(-1)} \lambda_{G_A}$. Recall, $\lambda_{G_A} = DFT.(e_0^T.G_A)^T$ (i.e. the DFT of the first row of the Gram matrix of the set A is the vector of eigenvalues of G_A). The exact form of M^{-1} was given in expression (A.17), and we can see that

$$e_0^T M^{(-1)} = \frac{1}{N} [1, -1, 1, \dots, 1, -1].$$

Thus, it holds that

$$p_0 = e_0^T M^{(-1)} DFT.(e_0^T.G_A)^T = \frac{1}{N} [1, -1, 1, \dots, 1, -1].DFT.(e_0^T.G_A).$$

We can see that that

$$[1, -1, 1, \dots, 1, -1].DFT = Ne_{N/2},$$

as this is equivalent to adding a π phase to each of the rows of the DFT matrix and then summing up the rows. Without the phase shift, the sum of the rows is a vector with a non-zero entry only at the first position. The phase shift corresponds to a cyclic permutation of columns by $N/2 - 1$ positions, so the sum of the rows of the permuted DFT matrix has the only non-zero entry at the $(N/2 + 1)^{st}$, and this entry is N . Hence we have

$$p_0 = e_{N/2}.(e_0^T.G_A)^T = \exp(-2\alpha^2),$$

and we have proven our Lemma. \square

- [1] Duan L M and Guo G C 1998 *Phys. Rev. Lett.* **80** 4999
- [2] Ferreyrol F, Barbieri M, Blandino R, Fossier S, Tualle-Broui R and Grangier P 2010 *Phys. Rev. Lett.* **104** 123603
- [3] Xiang G Y, Ralph T C, Lund A P, Walk N and Pryde G J 2010 *Nature Photonics* **4** 316
- [4] Zavatta A, Fiurasek J and Bellini M, 2011 *Nature Photonics* **5** 52
- [5] Menzies D and Croke S 2009 arXiv:0903.4181
- [6] Jeffers J 2011 *Phys. Rev. A* **83** 053818
- [7] Ivanovic I D 1987 *Phys. Lett. A* **123** 257
- [8] Dieks D 1988 *Phys. Lett. A* **126** 303
- [9] Peres A 1988 *Phys. Lett. A* **128** 19
- [10] Chefles A and Barnett S M 1998 *Phys. Lett. A* **250** 223
- [11] Chefles A and Barnett S M 1998 *J. Phys. A* **31** 10097
- [12] Chefles A, Jozsa R, and Winter A 2003 arXiv:quant-ph/0307227
- [13] Chefles A 2002 *Phys. Rev. A* **65** 052314
- [14] C. W. Helstrom *Quantum detection and estimation theory*, (Academic Press, New York, 1976)
- [15] Eldar Y C, Megretski A, and Verghese G C 2004 *IEEE Transactions on Information Theory* **50** 1198
- [16] Bennett C and Brassard G, *Proceedings of IEEE International Conference on Computers Systems and Signal Processing* (Bangalore, India, 1984) pp. 175–179
- [17] Barbosa G A, Corndorf E, Kumar P and Yuen H P 2003 *Phys. Rev. Lett* **90** 227901
- [18] Sych D and Leuchs G 2010 *New J. Phys.* **12** 053019

- [19] Broadbent A, Fitzsimons J, and Kashefi E *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Los Alamitos, USA, 2009) pp. 517–526
- [20] Andersson E, Curty M , and Jex I 2006 *Phys. Rev. A* **74** 022304
- [21] *Matrix Computations* 3rd ed. (The Johns Hopkins University Press, Baltimore, USA, 1996)
- [22] Pegg D T, Phillips L S, and Barnett S M 1998 *Phys. Rev. Lett.* **81** 1604
- [23] Babichev S A, Ries J, and Lvovsky A I 2003 *Europhys. Lett.* **64** 1
- [24] Dunjko V, Kashefi E and Leverrier A 2012 *Phys. Rev. Lett.* **108** 200502